

## Clustering of Absorbers

S. Cristiani<sup>1</sup>, S. D’Odorico<sup>2</sup>, V. D’Odorico<sup>3</sup>, A. Fontana<sup>4</sup>, E. Giallongo<sup>4</sup>, L. Moscardini<sup>1</sup>, S. Savaglio<sup>2</sup>

<sup>1</sup>*Dipartimento di Astronomia dell’Università di Padova, Vicolo dell’Osservatorio 5, I-35122 Padova, Italy*

<sup>2</sup>*European Southern Observatory, K.Schwarzschild-Strasse 2, D-85748 Garching bei München, Germany*

<sup>3</sup>*International School for Advanced Studies, SISSA, via Beirut 2-4, I-34014 Trieste, Italy*

<sup>4</sup>*Osservatorio Astronomico di Roma, via dell’Osservatorio, I-00040 Monte Porzio, Italy*

**Abstract.** The observed clustering of Lyman- $\alpha$  lines is reviewed and compared with the clustering of CIV systems. We argue that a continuity of properties exists between Lyman- $\alpha$  and metal systems and show that the small-scale clustering of the absorbers is consistent with a scenario of gravitationally induced correlations. At large scales statistically significant over and under-densities (including voids) are found on scales of tens of Mpc.

## 1 Introduction

The search for clustering of the Lyman- $\alpha$  lines has produced along the years diverse results. Systematic studies of the distribution of redshifts in the QSO Lyman- $\alpha$  forest began in the early 80’s with the work by Sargent et al. [34] who concluded that no structures could be identified. Almost all the subsequent results have failed to detect any significant correlation on velocity scales  $300 < \Delta v < 30000 \text{ km s}^{-1}$  [35, 2, 45]. On smaller scales ( $\Delta v = 50 - 300 \text{ km s}^{-1}$ ) there have been indications of weak clustering [44, 31, 7, 9, 25], together with relevant non-detections [30, 41, 21].

On the contrary, metal-line systems selected by means of the CIV doublet have been early recognized to show strong clustering up to  $600 \text{ km s}^{-1}$  [36], suggesting a different spatial distribution. In fact, the absence of power in the two-point correlation function has been claimed as a striking characteristic of the Lyman- $\alpha$  forest and has been used as a basic argument to develop a scenario of the Lyman- $\alpha$  absorbers as a totally distinct population with respect to metal systems and therefore galaxies.

## 2 The Lyman- $\alpha$ database

Our analysis of the clustering of the Lyman- $\alpha$  lines is based on data obtained in the framework of an ESO key-programme devoted to the study of QSO ab-

sorption systems at high redshifts. We have obtained spectra of several QSOs, with emission redshifts ranging from 3.27 to 4.12, at resolution between 9 and 14 km s<sup>-1</sup> [10]. We have complemented our data with other spectra available in the literature with similar resolution and redshift range, [6, 32, 33, 21, 20] obtaining a final sample of 14 QSOs (in the following *extended sample*), more than 1600 Lyman- $\alpha$  's, a suitable database to investigate clustering, especially at high redshift, where the density of lines provides particular sensitivity.

### 3 Small-scale clustering of Lyman- $\alpha$ lines

To study the clustering properties of our sample of Lyman- $\alpha$  lines we have adopted the two-point correlation function (TPCF) in the velocity space:

$$\xi(\Delta v) = \frac{N_{obs}(\Delta v)}{N_{exp}(\Delta v)} - 1 \quad (1)$$

In our case sample  $N_{exp}$ , the number of pairs expected in a given velocity bin from a random distribution in redshift, is obtained averaging 1000 numerical simulations of the observed number of redshifts, trying to account for all the relevant cosmological and observational effects. In particular the set of redshifts is randomly generated in the same redshift interval as the data according to the cosmological distribution  $\propto (1+z)^\gamma$ , where the best value of  $\gamma = 2.65$  has been derived from a maximum likelihood analysis [18]. The results are not sensitive to the value of  $\gamma$  adopted. Incomplete wavelength coverage due to gaps in the spectrum or line blanketing of weak lines due to strong complexes is also accounted for. Lines with too small velocity splittings, compared with the finite resolution or the intrinsic blending are excluded in the estimate of  $N_{exp}$ . The resulting correlation function for the full *extended sample* of Lyman- $\alpha$  lines shows a weak but significant signal, with  $\xi \simeq 0.2$  in the 100 km s<sup>-1</sup> bin: 739 pairs are observed while only 624 are expected for a random distribution, a 4.6 $\sigma$  deviation from poissonianity.

We have then explored the variations of the clustering as a function of the column density. For lines with  $\log N_{HI} \leq 13.6$  no evidence for clustering is present. On the contrary, for lines with  $\log N_{HI} \gtrsim 14$  (Fig. 1), the correlation function at  $\Delta v = 100$  km s<sup>-1</sup> shows a remarkable increase in amplitude ( $\xi \simeq 0.7$ ) and significance: 234 pairs are observed while only 145 are expected for a random distribution, a more than 7 $\sigma$  deviation from poissonianity. No relevant feature other than the peak at small velocity separations is observed. In particular previous claims for anti-clustering on scales  $\sim 600 - 1000$  km s<sup>-1</sup> are not confirmed. Fig. 2 shows more in detail the variation of the amplitude of the two-point correlation as a function of the column density. A trend of increasing amplitude with increasing column density is apparent.

We have also studied the evolution of the TPCF with the redshift for the sub-sample of Lyman- $\alpha$  lines with column densities  $\log(N_{HI}) > 13.8$ . The amplitude of the correlation at 100 km s<sup>-1</sup> decreases with increasing redshift

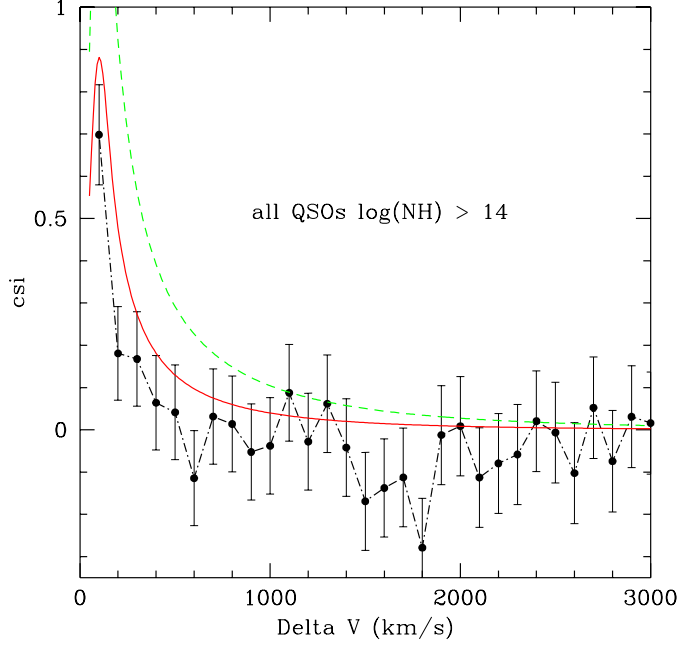


Figure 1: Two-point correlation function in the velocity space for lines with column densities  $> 10^{14} \text{ cm}^{-2}$ . Errors estimated with bootstrap resampling. The continuous line corresponds to a CDM model normalized to reproduce the observed abundance of clusters, the dashed line corresponds to a tilted CDM (see section 5).

from  $0.85 \pm 0.14$  at  $1.7 < z < 3.1$ , to  $0.74 \pm 0.14$  at  $3.1 < z < 3.7$  and  $0.21 \pm 0.14$  at  $3.7 < z < 4.0$ . Unfortunately, HST data are still at too low-resolution or are too scanty [1, 4] to allow a meaningful comparison with the present data. Nonetheless, the result by Ulmer [42], who measured with FOS data a  $\xi = 1.8^{+1.6}_{-1.2}$ , on scales of  $200 - 500 \text{ km/s}$  at  $0 < z < 1.3$ , suggests that the trend persists at lower redshifts.

#### 4 Comparison with CIV clustering and discussion

A discussion of the clustering of the CIV systems can be found in the paper by V.D’Odorico (this conference). A correlation of the clustering amplitude with column density is observed also for CIV systems. In the upper-right of Fig. 2 the TPCF derived for high and low column density CIV metal systems is shown. An extrapolation of the increasing amplitude trend observed for the TPCF of the Lyman- $\alpha$  lines would easily intercept the corresponding estimates

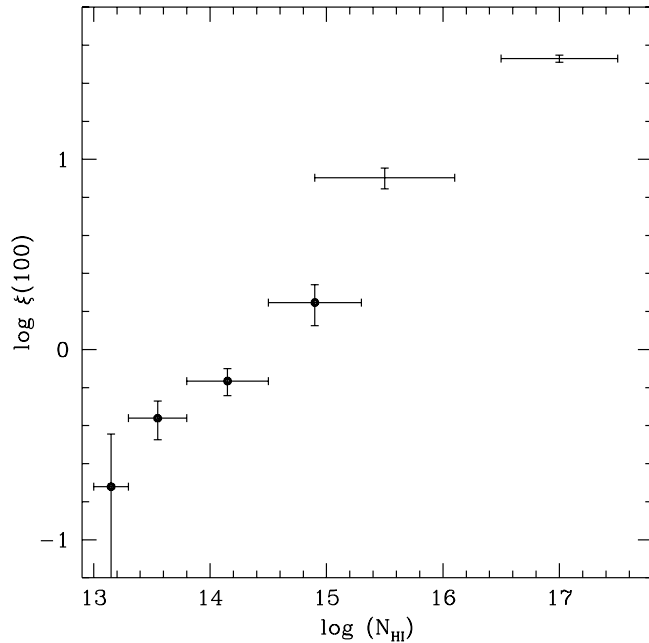


Figure 2: Variation of the amplitude of the two-point correlation function as a function of the column density threshold for the sample of the Lyman- $\alpha$  lines (filled circles). The two points in the upper-right side of the picture show the correlation of the CIV metal systems (see text)

derived from the CIV metal systems. The similarity between the shapes of the TPCF's of the Lyman- $\alpha$  and CIV systems, when observed at comparable resolution [29, 40], is also remarkable.

Fernandez-Soto et al. [17] have investigated the clustering properties of the Lyman- $\alpha$  clouds on the basis of the corresponding CIV absorptions, suggesting that CIV may resolve better the small-scale velocity structure that cannot be fully traced by Lyman- $\alpha$  lines. As a consequence, the estimates of the TPCF of the Lyman- $\alpha$  absorbers, although the effects of the non-negligible width of the lines (the “line-blanketing”) are considered in our simulations, should be regarded as a lower limit to the real clustering amplitude. However, it is not straightforward to translate the properties of the CIV absorbers in the corresponding ones of the Lyman- $\alpha$  absorbers, since observations show that the velocity structures of high and low-ionization species are often different. In any case, since the underestimation of the Lyman- $\alpha$  TPCF would be more severe at larger column densities, the trend of an increasing correlation with increasing column density appears to be real, from the lowest column densities up to those corresponding to the strongest metal systems.

If we add to this observation the following pieces of evidence:

1. At low redshift a considerable fraction of the Lyman- $\alpha$  lines has been observed to be associated with luminous galaxies and the local large-scale structure [22, 23];
2. Metallicities of the order  $10^{-2}$ , i.e. similar to the ones derived for the heavy elements absorptions originated in galactic halos, are observed for Lyman- $\alpha$  clouds with  $\log N_{HI} > 14$  [8];
3. The volume density and cosmological evolution of the same  $\log N_{HI} > 14$  clouds are similar to those of the damped systems [18];

all together this suggests a physical association between the Lyman- $\alpha$  clouds with  $\log N_{HI} > 14$  and the halos of protogalactic systems.

The typical column density below which no clustering of the Lyman- $\alpha$  lines is observed ( $\log(N_{HI}) \sim 13.6$ ) corresponds to the position of the break in their column density distribution [28, 18], which has been identified with the transition from a variety of systems in various stages of gravitational infall and collapse (or even under-densities) to gas associated with star forming galaxies [47, 26].

The observed clustering properties are qualitatively consistent with a scenario of gravitationally induced correlations. The dependence of the clustering amplitude on the column density arises naturally in models involving biasing for the formation of structures in the universe: objects associated with the stronger potential wells are expected to be more clustered and the HI column density is directly related to the depth of the well (and to the associated mass).

The trend of increasing correlation with decreasing redshift is also a strong prediction of any model of structure formation based on gravitational instability. On the contrary, theories of explosive structure formation [43, 46] expect a velocity correlation either unchanging or diminishing with time.

To be more quantitative an estimate of the masses of the clouds is needed. Known the ionization field from the proximity effect [18] and the sizes of the clouds from quasar pairs and gravitational lenses ( $\sim 100h^{-1}$  Kpc) [38, 39, 3, 12, 15], the observed column densities can be transformed in masses with a photoionization code [16]. Then, we can apply a recipe connecting a given model of Universe, the spectrum of the initial fluctuations, to the clustering of the mass [24] and, finally, convolve it with the effects of the non-negligible size of the absorbing structures and their peculiar velocities  $\sigma_v$  [19]. The results for two standard cosmological models, a CDM normalized to the abundance of clusters and a tilted CDM, are shown in Fig. 1 for  $\sigma_v \sim 50$  km/s.

## 5 Large-scale Clustering

## 5.1 Voids

A typical way of looking for non-random fluctuations of the line density on large scales is the search for voids. Voids in the Lyman- $\alpha$  forest provide a test for models of the large-scale structure in the Universe and of the homogeneity of the UV ionizing flux. Previous searches for megaparsec-sized voids have produced a few claims [11, 14], but uncertainties in the line statistics strongly influence the probability estimate [27]. High resolution data, less affected by blending effects, are ideal also for the study of voids. Detections of voids in individual cases are interesting but a more general approach, assessing more quantitatively how common is the phenomenon and avoiding the pitfalls of “a posteriori statistics”, is preferable. We have adopted the following procedure: for each of the objects in the extended sample and for lines with  $\log(N_{HI}) \geq 13.3$  we have estimated through Montecarlo simulations the void size  $\Delta r$  (in comoving coordinates) for which the probability to find at least one void  $\geq \Delta r$  is 0.05. Then we have searched each spectrum for voids of dimensions equal or larger than  $\Delta r$ . In 3 out of the 9 cases for which the density of lines and the absence of spectral gaps allow a meaningful analysis, at least one void was detected (two voids for 0055 – 26). The binomial probability corresponding to such occurrences is  $8 \cdot 10^{-3}$ .

From the spectra it is apparent that the regions corresponding to the voids are not completely devoid of lines: weak absorptions are observed within the voids. This agrees with low-redshift observations [37] that have shown that in the local Universe voids are not entirely devoid of matter.

Even if underdense regions are statistically significant in our sample, the filling factor is rather low: the voids cover only about 2% of the available line-of-sight path-length, confirming previous results [5].

## 5.2 Over- and under-densities of lines

Voids are just an extreme case of spectral regions showing an under-density of lines. Various theoretical reasons prompt to tackle the issue of over- and under-densities of lines from a more general point of view: the typical signature of a “proximity effect” due to a foreground quasar is a lack of weak lines in the Lyman forest, rather than lack of lines in general [13]; the relative filling factor of under-densities and over-densities may provide a constraint on the theories of structure formation.

We have analysed the spectra of the extended sample with a counts-in-cells technique searching for over- and under-densities of lines with  $\log(N_{HI}) \geq 13.3$  on scales from 10 to 80 Mpc and comparing the observed counts with Montecarlo simulations in order to assess the significance of the deviations. On smaller scales the shot noise is too large, on larger scales the “integral constraint”, forcing the simulated number of lines to be equal to the observed one, prevents from the possibility of detecting any deviation. The threshold to define significant a deviation (for example in excess) in a given spectral

interval has been set in a way that at a given scale for a given quasar there is a 0.05 probability of observing at least one deviation of this type on the whole spectrum for a locally Poissonian distribution.

5 QSOs out of 15 show at least one over-density in their spectrum at 10 Mpc scales, corresponding to a binomial probability of  $6 \cdot 10^{-4}$  of being drawn from a poissonian distribution of lines. 4 QSOs show at least one over-density of lines at 20 and 30 Mpc scales, corresponding to a binomial probability of  $5 \cdot 10^{-3}$ . At larger scales the number of significant over-densities decreases. None is observed at 80 Mpc.

Under-densities appear to be roughly as common as over-densities, once the lack of sensitivity at lower redshifts and smaller scales, due to the low density of lines, is taken into account.

The existence of a roughly equal number of over and under-densities on scales 10 – 40 Mpc is easily understandable in terms of the linear theory of the evolution of the perturbations, that is a plausible approximation at such relatively high redshifts: gravity has not yet had time to give a significant skewness to the (under)over-density distribution. Besides, almost any hierarchical clustering scenario would expect that at  $z \sim 2 - 4$  gravity has not yet had enough time to transfer power on 80 Mpc scales and give origin to significant over or under-densities.

## References

- [1] Bahcall J., Bergeron J., Boksenberg A., et al., 1996, ApJ 457, 19
- [2] Bechtold J., 1987, in J. Bergeron et al. eds., Proc. Third IAP Workshop, High Redshift and Primeval Galaxies. Editions Frontieres, Gif-sur-Yvette, p. 397
- [3] Bechtold J., Crotts A. P. S., Duncan R. C., Fang Y., 1994, ApJ 437, L83
- [4] Brandt J. C., Heap S. R., Beaver E. A., et al., 1995, AJ 105, 831
- [5] Carswell R. F., Rees M. J., 1987, MNRAS 224, 13p
- [6] Carswell R.F., Lanzetta K.M., Parnell H.C., Webb J.K., 1991, ApJ 371, 36
- [7] Chernomordik V.V., 1995, ApJ 440, 431
- [8] Cowie L.L., Songaila A., Kim T., Hu E.M., 1995, AJ 109, 1522
- [9] Cristiani S., D’Odorico S., Fontana A., Giallongo E., Savaglio S., 1995, MNRAS 273, 1016
- [10] Cristiani S., D’Odorico S., D’Odorico V., et al., 1997, MNRAS 285, 209
- [11] Crotts A. P. S. 1989, ApJ 336, 550
- [12] Dinshaw N., Foltz C.B., Impey C.D., Weymann R.J., Morris S.L. 1995, Nature 373, 223
- [13] Dobrzycki, A., Bechtold, J. 1991a, in Crampton ed., Proc. Workshop The space distribution of quasars, ASP Conference series 21, p. 272
- [14] Dobrzycki, A., Bechtold, J. 1991b, ApJ 377, L69
- [15] D’Odorico S., et al., this conference
- [16] Ferland G.J., 1996, Hazy, a Brief Introduction to Cloudy, University of Kentucky Department of Physics and Astronomy Internal Report
- [17] Fernandez-Soto A., Lanzetta K.M., Barcons X., Carswell R.F., Webb J.K., Yahil A., ApJ 460, L85

- [18] Giallongo E., Cristiani S., D’Odorico S., Fontana A., Savaglio S., 1996, ApJ 466, 46
- [19] Heisler, J., Hogan, C. J., White, S. D. M. 1989, ApJ 347, 52
- [20] Hu E. M., Kim T., Cowie L. L., Songaila A., Rauch M., 1995, AJ 110, 1526
- [21] Kirkman D., Tytler D., 1997 astro-ph/9701209
- [22] Lanzetta K. M., Bowen D. B., Tytler D., Webb J. K., 1995, ApJ 442, 538
- [23] Le Brun V., Bergeron J., Boisse P., 1996, A&A 306, 691
- [24] Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, MNRAS 286, 115
- [25] Meiksin A., Bouchet F. R., 1995, ApJ 448, L85
- [26] Mücke J. P., Petitjean P., Kates R. E., Riediger R., 1996, A&A 308, 17
- [27] Ostriker J. P., Bajtlik S., Duncan R. C., 1988, ApJ 327, L35
- [28] Petitjean P., Webb J.K., Rauch M., Carswell R.F., Lanzetta K., 1993, MNRAS 262, 499
- [29] Petitjean P., Bergeron J., 1994, A&A 283, 759
- [30] Pettini M., Hunstead R. W., Smith L. J., Mar D. P. 1990, MNRAS 246, 545
- [31] Rauch M., Carswell R. F., Chaffee F. H., Foltz C. B., Webb J. K., Weymann R. J., Bechtold J., Green R. F. 1992, ApJ 390, 387
- [32] Rauch M., Carswell R. F., Webb J. K., Weymann R.J., 1993, MNRAS 260, 589
- [33] Rodríguez-Pascual P.M., De La Fuente A., Sanz J.L., Recondo M.C., Clavel J., Santos-Lleó M., Wamsteker W., 1995, ApJ 448, 575
- [34] Sargent W. L. W., Young P. J., Boksenberg A., Tytler D. 1980, ApJS 42, 41
- [35] Sargent W. L. W., Young P. J., Schneider D. P., 1982, ApJ 256, 374
- [36] Sargent W. L. W., Boksenberg A., Steidel C. C. 1988, ApJS 68, 539
- [37] Shull J.M., Stocke J.T., Penton S., 1996, AJ 111, 72
- [38] Smette A., Surdej J., Shaver P.A. et al., 1992, ApJ 389, 39
- [39] Smette A., Robertson J.G., Shaver P.A., Reimers D., Wisotzki L., Köhler Th., 1995, A&AS 113, 199
- [40] Songaila A., Cowie L.L., 1996, AJ 112, 335
- [41] Stengler-Larrea E. A., Webb J. K. 1993, in Chincarini G., Iovino A., Maccacaro T., Maccagni D. eds., *Observational Cosmology*, ASP Conference Series, 51, 591
- [42] Ulmer A., 1996, ApJ 473, 110
- [43] Vishniac E.T., Ostriker J.P., Bertschinger E., 1985, ApJ 291, 399
- [44] Webb J. K., 1987, in Hewett A., Burbidge G., Fang L. Z. eds., *Proc. IAU Symp.* 124, *Observational Cosmology*. Reidel, Dordrecht, p. 803
- [45] Webb J. K., Barcons X., 1991, MNRAS 250, 270
- [46] Weinberg D.M., Ostriker J.P., Dekel A., 1989, ApJ 336, 9
- [47] Zhang Y., Anninos P., Norman M.L., 1996, ApJ 459, 12